

Python For Data Science Course

Introduction

This course is designed to empower the students to understand the fundamentals of machine learning, their application areas, and how machine learning algorithms work and provide hands-on interactive learning with guided machine learning projects. The estimated course duration will take around 2 months to complete the course. Assignments will be given to the students weekly so that they can practice more of what they have learned from their course.

In the hands-on exercise, the students get a chance to understand the problem statement, scope the problem, and perform data cleaning, analysis, and modeling from scratch.

After learning the course the students are expected to:

- 1) Understand the importance of data and how it can transform the business.
- 2) Code in Python and can write efficient codes for manipulating data.
- 3) Learn the following tools and technologies:
 - a) Scikit learn
 - b) Numpy
 - c) Pandas
 - d) Data Visualization
 - e) Jupyter Notebooks
 - f) Python
- 4) Can easily start working on any problems related to classification & regression.

Table of Contents

Python Fundamentals

1. Getting Started with Python
2. Install anaconda and Python on Windows and Linux
3. Variables and Data Types (String, Integer, Float, List, Tuples, Dictionary)
4. Lists and For Loops
5. Conditional Statements
6. Dictionaries and Frequency Tables
7. Functions and their types
8. Concept of modules
9. Working with OS
10. Command Line Tool
11. Working with files
12. Argument Parser In Python

- 13. Working with Datetime
- 14. Assignments

Object Oriented Programming In Python

- 15. Introduction To OOPS In Python
- 16. Class and Objects
- 17. Constructor In Python
- 18. Inheritance and its types
- 19. Polymorphism(Method Overloading & Operator Overloading)
- 20. Methods and Type of methods(Instance, Class & Static Methods)
- 21. Creating a class to load csv data
- 22. Assignments Related to Class

Getting Started With Data Science (Numpy, Pandas & Notebooks)

- 23. Python For Data Science
- 24. Working with Jupyter Notebook
- 25. Introduction To Numpy & Pandas
- 26. Some Useful Numpy & Pandas Operations
- 27. Pandas GroupBy Function, Aggregate Function

Data Science Fundamentals(Data Cleaning & EDA)

- 28. Data Cleaning
- 29. Exploratory Data Analysis
- 30. Univariate Analysis
- 31. Bivariate Analysis
- 32. Data Visualization using matplotlib, Plotly or seaborn

Understanding Machine Learning Algorithms

- 33. Intro To Machine Learning
- 34. Types Of Learning
- 35. Supervised Learning Algorithms(Classification & Regression)
- 36. Classification: Logistic Regression, Decision Trees, Random Forest, KNN
- 37. Evaluation Metrics For Classification Problems
- 38. Regression: Linear Regression
- 39. Evaluation Metrics For Regression Problem
- 40. Unsupervised Learning: Clustering

Hands-On Machine Learning Practice Problem(Depending upon course timeline)

41. Guided Project On Classification (Crop Health Prediction, Heart Stroke Prediction, Accident Severity Prediction, Employee Promotion Classification) (any one)
42. Guided Project On Regression (Car Price Prediction, House Price Prediction, Avocado Price Prediction, Ticket Fare Prediction) (anyone)
43. Final Assignment For Students : Any project on Regression and Classification

Estimated Timeline For the course = 2.5 months